Scale-up of an RNA reference standard for high-throughput microarray QC

Paul K. Wolber, Ph.D.
Project Manager
28-Mar-2003

(with thanks to a host of others
 at Agilent and Rosetta…)

Agilent Technologies

Problem Statement

- **Goal & Constraints**

  + **Sample for final product lot QC (hybridization assay)**

    - **Fixed QC array design**

    - **Sensitivity to known & potential error modes**

  + **Sample must be**

    - **Reproducible**

    - **Manufacturable**

    - **Economical**

Possibilities Considered

- **Oligo-only sample**

- **Complex natural sample**

- **Complex synthetic sample**

Possibilities Considered: Details

- **Oligo-only sample**

  + **Easiest to make & maintain**

  + **Limited relevance to customer experience**

    - **Narrow dynamic range**

    - **No cRNA component**

    - **Difficult to generate expression ratio data**

    i. **Need to develop a family of oligos**

    ii. **Need to develop QC and formulation methods to maintain ratios**

Possibilities Considered: Details

- **Complex natural sample**

  + **Hardest to make and maintain**
    - **More labor to make than to use**
    - **Qualifying new batches of natural mRNA difficult**
    - **Species-specific**

  + **Most relevant to customer experience**
    - **Full dynamic range**
    - **Labeled cRNA, as used by customers**
    - **If 2 natural samples are used, they can be chosen to generate a rich range of differential expression (but TRUTH difficult to determine…)**
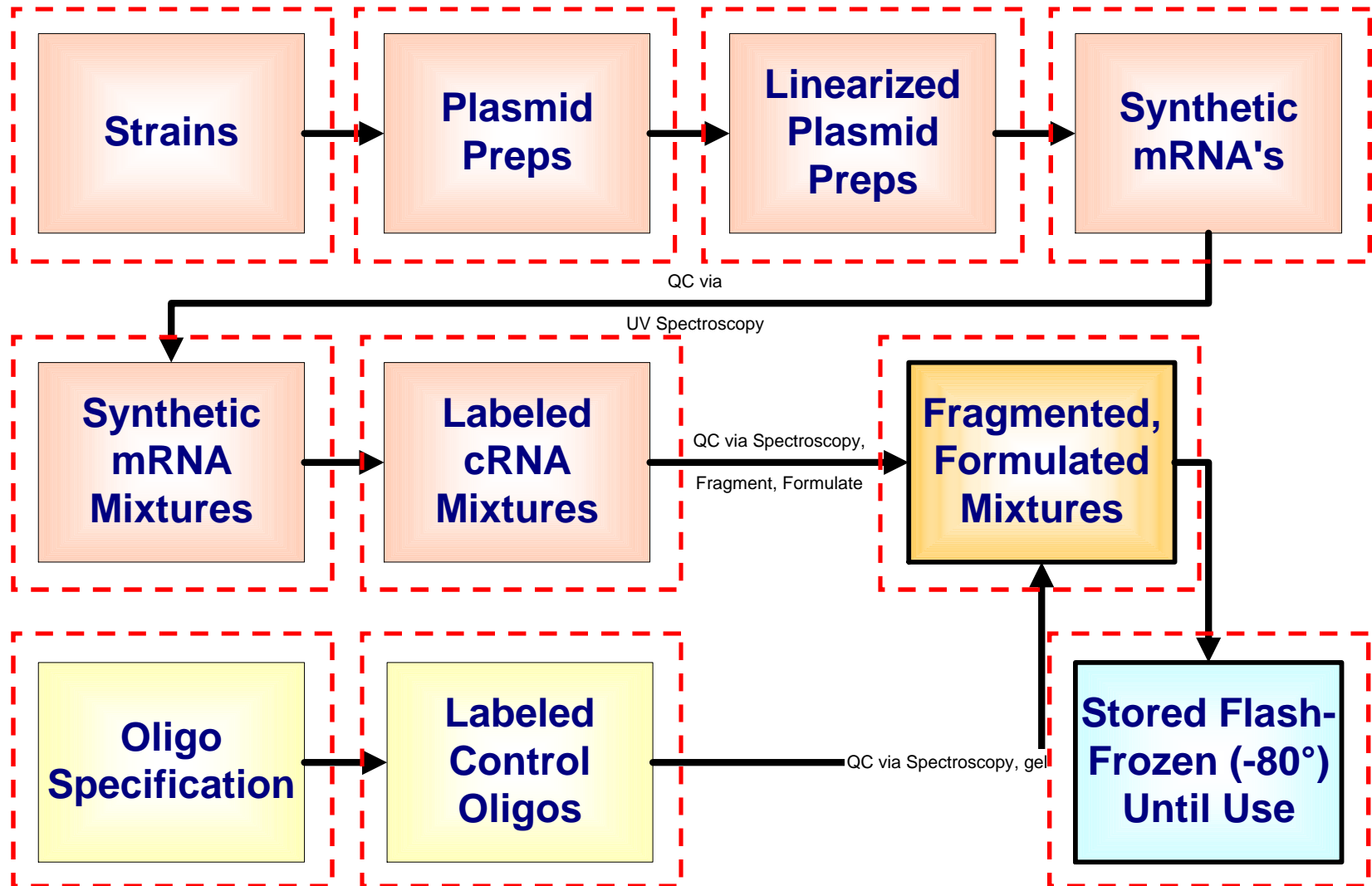
Agilent Technologies

Possibilities Considered: Details

✓ • **Complex synthetic sample (oligo + E1A cRNA)**

+ **Intermediate difficulty of production/maintenance**

  - **300x more efficient than "natural" sample**

  - **Stable source of (synthetic) mRNA component**

  - **Species-independent**

+ **Intermediate relevance to customer experience**

  - **~200-fold dynamic range**

  - **Labeled cRNA, as used by customers**

  - **Known ratios of E1A targets in 2 samples**

  - **Control relative specific activities by labeling mixture**

Biomaterials Flow

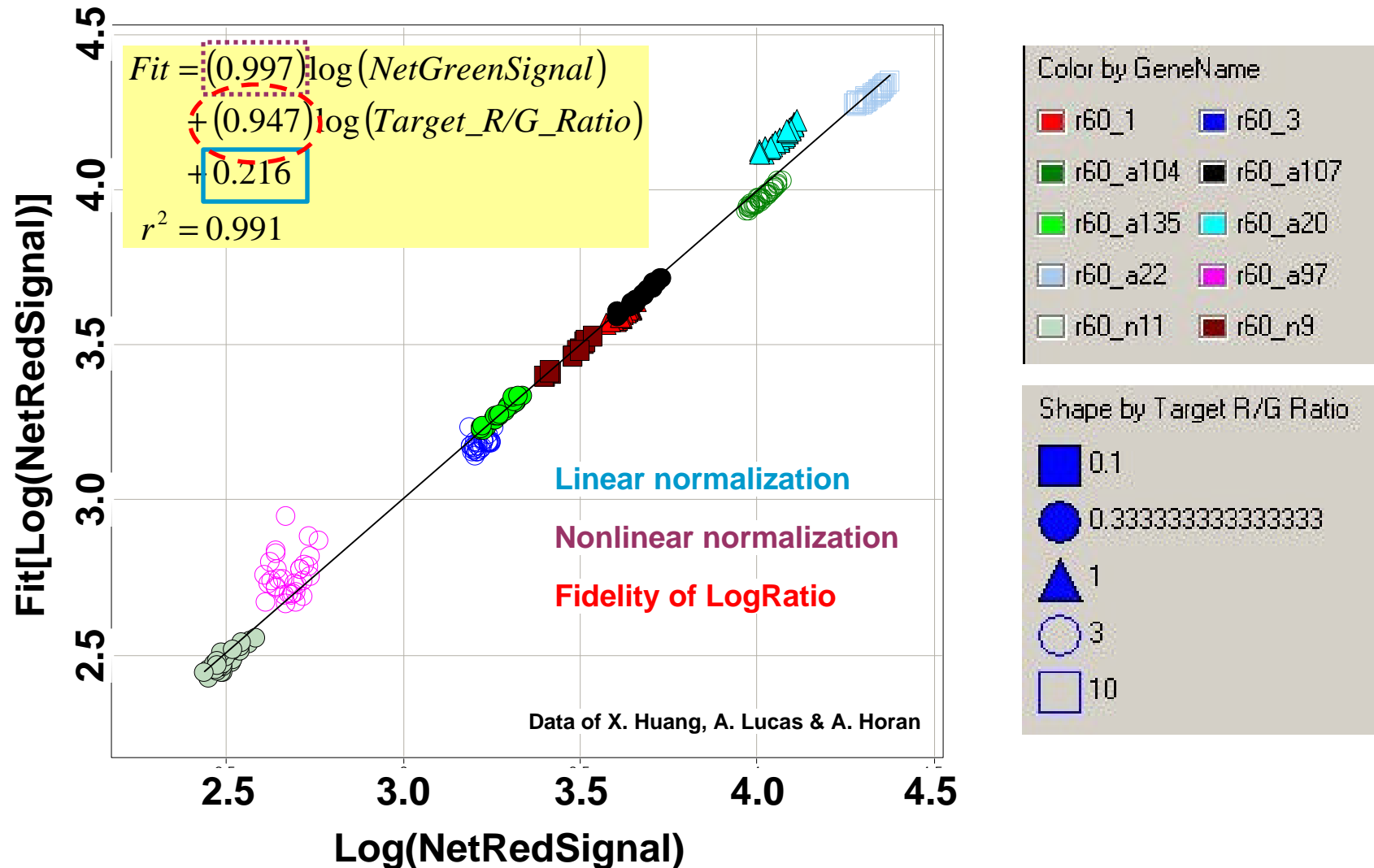Strains → Plasmid Preps → Linearized Plasmid Preps → Synthetic mRNA's

QC via UV Spectroscopy

Synthetic mRNA Mixtures → Labeled cRNA Mixtures → QC via Spectroscopy, Fragment, Formulate → Fragmented, Formulated Mixtures → Stored Flash-Frozen (-80°) Until Use

Oligo Specification → Labeled Control Oligos → QC via Spectroscopy, gel

Typical E1A Cocktail Compositions*

Transcript	Nominal Copies per Cell		Target Ratio
	Cocktail 11	Cocktail 12	
r60_a20	100	100	1:1
r60_1	10	10	1:1
r60_a22	10	100	1:10
r60_n9	100	10	10:1
r60_a104	10	30	1:3
r60_a107	30	10	3:1
r60_3	3	9	1:3
r60_a135	9	3	3:1
r60_a97	0.5	1.5	1:3
r60_n11	1.5	0.5	3:1

* courtesy of Rosetta Inpharmatics

Multiple Regression Model of E1A Data

$$Fit = (0.997)\log(NetGreenSignal)$$
$$+ (0.947)\log(Target\_R/G\_Ratio)$$
$$+ 0.216$$
$$r^2 = 0.991$$

Linear normalization

Nonlinear normalization

Fidelity of LogRatio

Data of X. Huang, A. Lucas & A. Horan

Fit[Log(NetRedSignal)]

Log(NetRedSignal)

Color by GeneName

- r60_1
- r60_3
- r60_a104
- r60_a107
- r60_a135
- r60_a20
- r60_a22
- r60_a97
- r60_n11
- r60_n9

Shape by Target R/G Ratio

- 0.1
- 0.333333333333333
- 1
- 3
- 10

Fixed QC Array Design

Probes to oligo targets:

> **Probes to cRNA targets:**
>
> **Customer-oriented, ratio-centric measurements**

Measurement of known printing and synthesis error modes

Routine Quality Metrics Based on E1A Targets

- **Ratio Sensitivity**

  + **% of 1 copy/cell probes yielding ratios within 50% of expected value**
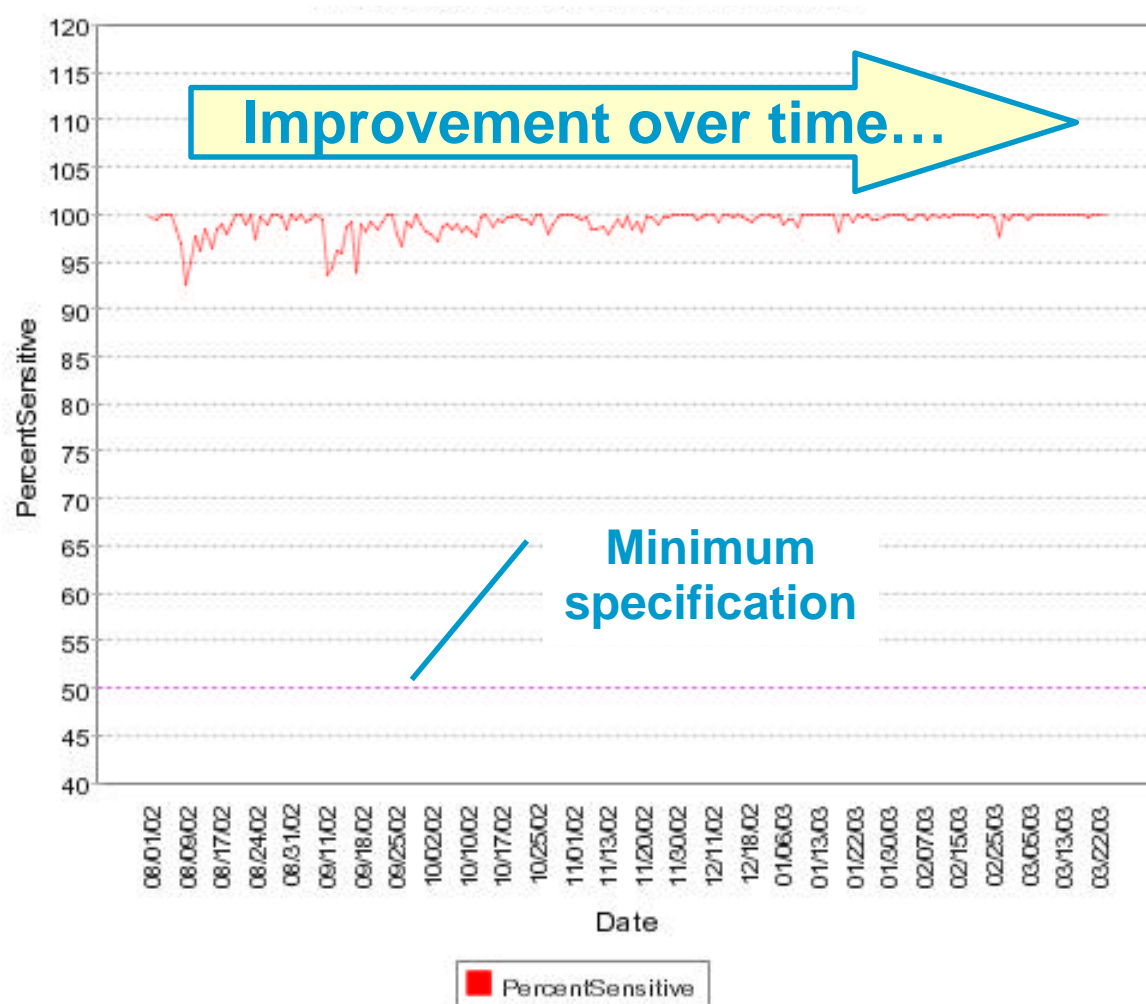
- **Reproducibility**

  + **Maximum value of $s_{LogRatio}$ for 5 different E1A probes**

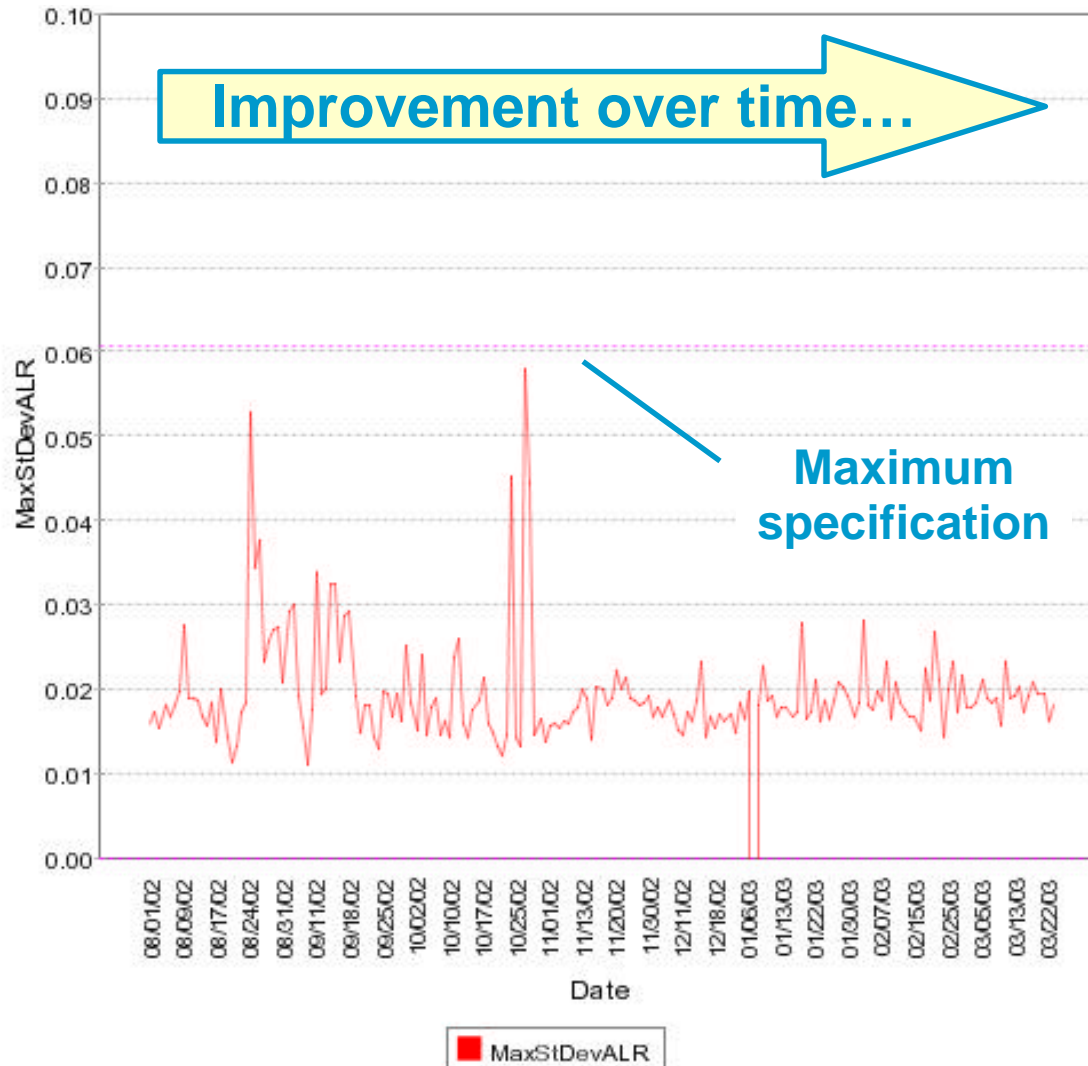 + **Indicates CV of LogRatio Measurements**

- **Accuracy**

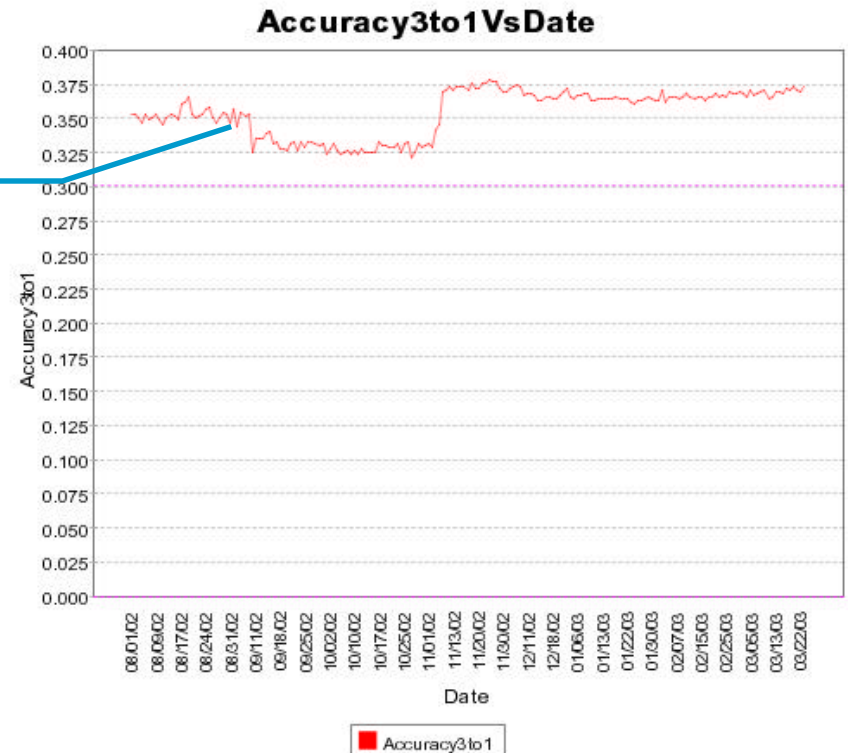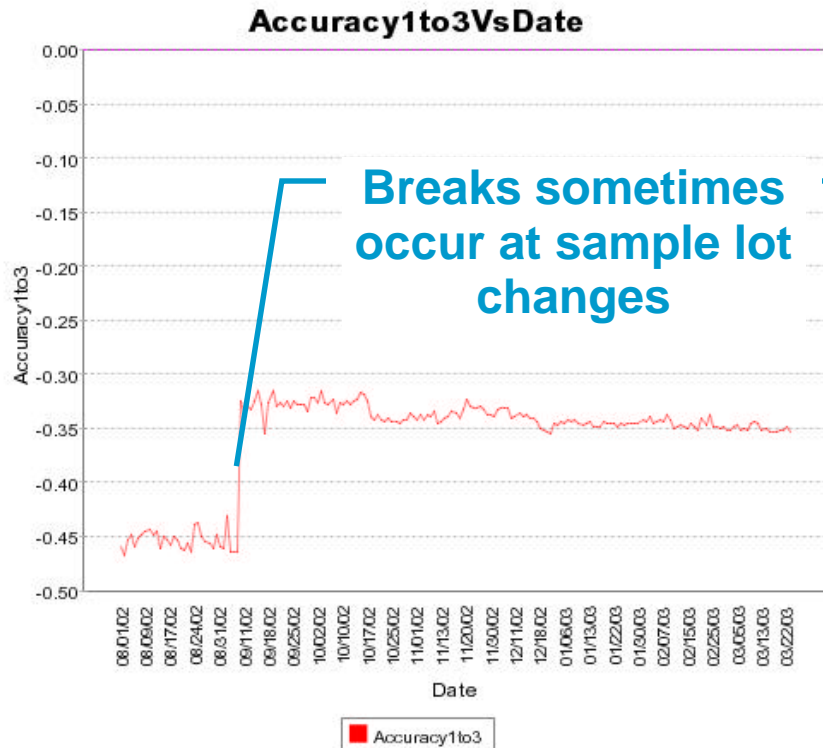 + **Average LogRatios (n=30) for 3:1 and 1:3 probes**

Run Chart: Ratio Sensitivity Metric

Run Chart: Reproducibility Metric

Run Charts: Accuracy Metrics

Breaks sometimes occur at sample lot changes

Conclusions

- **High-throughput micoarray QC presents some unique challenges**

  + **Long-term stability of standard sample**

  + **Cost-effectiveness of standard sample**

- **The Rosetta E1A spike-in system can be used to perform high-throughput microarray QC**

  + **Manufacturable (>10,000 assays to date)**

  + **Good compromise between simple and complex samples**

  + **Ratio-centric metrics**

Relevance to NIST Microarray Standards Initiative

- **Synthetic mRNA samples are a viable approach to standard generation and maintenance**

- **Rosetta E1A standard set is particularly attractive**

  + **Species-independent**

  + **Proven track record**

  + **Easily extended (by cloning additional inserts)**

  + **Suitable for use in multiple systems & settings**